

UNE APPROCHE MODÈLE MIXTE POUR LA CLASSIFICATION SUPERVISÉE DE SIGNAUX ÉLECTROPHYSIOLOGIQUES

Juliette Spinnato ^{1,2} & Marie-Christine Roubaud ¹ & Bruno Torr sani ¹
& Laurence Casini ² & Boris B rle ²

¹ LATP/ Aix-Marseille Universit , 39 rue F. Joliot Curie 13453 Marseille Cedex 13

² LNC / Aix-Marseille Universit , 3 place Victor-Hugo 13331 Marseille cedex 3

³ juliette.spinnato@etu.univ-provence.fr

R sum . La classification de donn es fortement bruit es et pr sentant une importante variabilit  constitue un enjeu important. Les signaux  lectrophysiologiques (EEG) correspondent   ce type de donn es complexes. Pour une t che cognitive  tudi e, l'activit  c r brale associ e pr sente une importante variabilit  temporelle d'une r p tition   une autre et le signal d'int r t, de faible amplitude, est *noy * dans un ensemble d'artefacts et de bruit. Nous proposons une proc dure permettant de classer, chez un m me sujet, deux types de signaux c r braux   partir d'une mod lisation par le mod le lin aire mixte gaussien. Ce travail s'inspire de la m thode pr sent e par Huang *et al.* (2008). Notre contribution r side d'une part dans une formalisation simplifi e de la mod lisation et d'autre part dans l'introduction d'une transformation en ondelettes permettant une r duction de la dimension temporelle sans perte d'information.

La proc dure pr sent e est appliqu e   la d tection d'ondes d'erreurs au cours d'une t che cognitive, et nous  tudions la performance de notre m thode sur six sujets en comparant les r sultats obtenus aux r sultats d'une analyse factorielle discriminante.

Mots-cl s : Classification supervis e, Mod le lin aire mixte, Transform e en ondelettes discr te, Signaux EEG, Variabilit  temporelle inter-essais, Onde d'erreur.

Abstract. The classification of noisy-data with high variability is an important issue. Electrophysiological (EEG) signals correspond to this type of complex data. In a Reaction-Time task, brain activity has an important temporal variability accross repetitions and the signal of interest, which is of low amplitude, is embedded into a set of artifacts and noise. For a fixed subject, we propose a procedure in order to classify two types of brain signals using a parametric modeling based on the Gaussian Mixed-Effects model. The present work is based on the method presented by Huang *et al.* (2008). Our contribution is first a simplified formalization of the modeling and second the introduction of a wavelet transform which leads to a reduction of temporal dimension without loss of information.

The procedure is applied to the detection of error-negativities during a cognitive task, and we study the performance of our method on six subjects by comparing the results with those obtained with a discriminant analysis.

Keywords. Supervised classification, Linear mixed-effects model, Discrete wavelets transform, EEG signals, Between-trials temporal variability, Error-negativity.

1 Introduction

Nous nous intéressons à la classification supervisée de deux classes de signaux fortement bruités et étudions l’apport de la modélisation pour ce type de problème. Plus particulièrement, nous proposons une procédure discriminante de modélisation pour les signaux électrophysiologiques (EEG).

Enregistrés au cours d’expériences cognitives et provenant de l’activité électrique du cerveau, ces signaux, de très faible intensité, sont parasités par la présence de nombreux artefacts (mouvements oculaires, clignement des yeux, ...) et d’un bruit important. Afin d’extraire l’activité électrique liée à la tâche que le sujet accomplit, il est d’usage de répéter cette même tâche plusieurs fois (les répétitions sont appelées *les essais*) et de baser l’analyse sur la moyenne des essais. Ce faisant, la variabilité inter-essais, qui contient de l’information utile, n’est pas exploitée.

La modélisation statistique est une solution proposée pour améliorer le traitement essai par essai. Plusieurs travaux ont été développés dans ce sens en utilisant le modèle mixte (e.g. Huang *et al.* (2008), Fazli *et al.* (2011)). Ce modèle a été également adapté à des données fonctionnelles avec l’utilisation d’une transformation en ondelettes discrète (DWT : Discrete Wavelet Transform)(e.g. Davidson (2009)). Ce travail s’inspire principalement de la démarche de Huang *et al.* (2008) qui ont modélisé la variabilité temporelle inter-essais par un modèle linéaire mixte gaussien (MLMG) afin de détecter des signaux EEG particuliers pour un sujet donné. Nous proposons d’une part une formalisation simplifiée de leur modélisation et d’autre part l’introduction d’une transformation en ondelettes permettant une réduction de la dimension temporelle sans perte d’information.

Dans cet article, l’ensemble de la procédure est détaillée en distinguant l’étape d’apprentissage et l’étape test et des résultats obtenus sur des données EEG réelles sont présentés.

2 Procédure

Les données sont structurées sous forme ”cubique”. Pour un sujet donné, les signaux EEG ont trois dimensions : le nombre d’essais N , le nombre d’électrodes M et le temps t_1, \dots, t_n . Ainsi, un essai brut i issu d’une classe c , noté B_i^c est sous la forme présentée en Figure 1.

A partir des données obtenues sur une session expérimentale, la méthode présentée se compose de deux étapes : l’apprentissage et le test. En phase apprentissage, pour chacune des classes, on tire aléatoirement un échantillon avec le même nombre d’essais. Sur les échantillons d’apprentissage, les essais de chaque classe sont modélisés par le MLMG et les paramètres des modèles sont estimés. En phase test, chaque nouvel essai est classé *via* un test du type rapport de vraisemblance construit à partir des estimations faites en phase d’apprentissage.

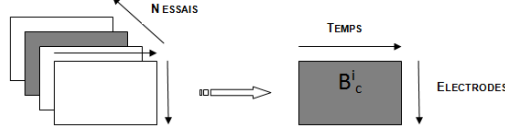


Figure 1: Des données cubiques de grandes dimensions

Réduction de dimension

Une étape de pré-traitement est proposée afin de réduire les dimensions des données brutes et d'en extraire les informations les plus pertinentes pour la classification.

La première réduction porte sur la dimension spatiale. Elle est définie en phase d'apprentissage. Une Analyse Factorielle Discriminante (AFD) entre les deux classes d'essais permet de définir les J combinaisons linéaires des M électrodes les plus discriminantes, appelées *les canaux*. Les mêmes combinaisons linéaires sont ensuite appliquées comme pré-traitement sur les données test.

La seconde réduction de dimension est temporelle. Elle se fait par une transformation en ondelettes discrète. Les signaux EEG étant de basse fréquence, la réduction de dimension temporelle se fait par une suppression des premiers niveaux de décomposition. Nous noterons n le nombre de coefficients d'ondelettes associés à chacun des J canaux après réduction.

Modélisation

Basée sur le MLMG, la modélisation proposée *décompose* les essais d'une même classe en deux parties : une partie fixe correspondant au comportement moyen de la classe à laquelle s'ajoute une partie aléatoire qui spécifie la variabilité d'un essai par rapport à la moyenne. La modélisation proposée est appliquée aux coefficients d'ondelettes $y(\lambda)$ obtenus par l'application d'une DWT au signal temporel de chaque essai (après réduction de dimension), où $\lambda \in \Lambda$ est un couple *(temps, échelle)* indexant les coefficients d'ondelettes.

Dans une classe donnée, les coefficients d'ondelettes $y_{ij}(\lambda)$ associés aux essais $i \in \{1, \dots, n\}$, et aux canaux $j \in \{1, \dots, J\}$ sont modélisés par :

$$y_{ij}(\lambda) = \mu_j(\lambda) + d_i(\lambda)b_i + \varepsilon_{ij}(\lambda)$$

où $\mu_j(\lambda)$ est le coefficient moyen de la classe correspondant au canal j et au couple λ , $d_i(\lambda)$ est le score discriminant associé à λ pour l'essai i .

Les effets aléatoires b_i des essais i sont des variables gaussiennes indépendantes centrées et de variance τ^2 et les résidus $\varepsilon_{ij}(\lambda)$ sont des variables gaussiennes indépendantes centrées de variance σ^2 et indépendantes des b_i .

Sous forme vectorielle le modèle s'écrit :

$$y_i = \mu + Z_i b_i + \varepsilon_i$$

où $y_i \in \mathbb{R}^{J \times n}$ est l'essai vectorisé i , $\mu \in \mathbb{R}^{J \times n}$ est la moyenne théorique des essais de la classe, $Z_i b_i$ est la partie aléatoire qui modélise la variabilité *pertinente* de l'essai i pour la discrimination et $\varepsilon_i \in \mathbb{R}^{n \times M}$ est le vecteur aléatoire des résidus, centré et de matrice de variance-covariance

$\sigma^2 I_{n \times M}$. Le vecteur Z_i est de la forme $Z_i = Z.d_i$ où $Z \in \mathcal{M}^{J \times n, n}$ est la matrice design associée aux canaux et $d_i \in \mathbb{R}^n$ est un vecteur de score de discrimination pour l'essai i .

Cette modélisation se base sur une approche *bayésienne empirique* : supposant connu *a priori* le vecteur des scores $d_i = (d_i(\lambda))'_{\lambda \in \Lambda} \in \mathbb{R}^n$, le modèle correspond à un MLMG pour les essais i d'une classe donnée et on a

$$y_i | d_i \sim \mathcal{N}(\mu, V_i)$$

où $V_i = \tau^2 Z d_i d_i' Z' + \sigma^2 I_{n \times M}$ est la matrice de variance-covariance de dimension $J \times n$.

Le calcul des vecteurs des scores est effectué à l'aide d'une AFD entre les deux classes. On détermine l'axe Δ_{v_1} qui discrimine au mieux les $y_i(\lambda) = (y_{i1}(\lambda), \dots, y_{iJ}(\lambda))'$ des deux classes dans la phase apprentissage et on prend pour score discriminant associé à $y_i(\lambda)$ sa coordonnée sur cet axe: $d_i(\lambda) = \langle v_1, y_i(\lambda) \rangle$.

Estimation

En phase d'apprentissage, pour chaque modèle c construit, les paramètres inconnus μ^c, τ_c^2 et σ_c^2 sont estimés via les procédures classiques d'estimation du MLMG.

En phase test, les paramètres du modèle sont remplacés par leur estimation (plug-in). Pour chaque nouvel essai y_{itest} le vecteur score est calculé et l'effet aléatoire b_{itest}^c est prédit dans chaque classe c par

$$\hat{b}_{itest}^c = \frac{1}{\hat{\sigma}^2} (Z d_{itest})' (\hat{V}_{itest}^c)^{-1} (y_{itest} - \hat{\mu}^c)$$

où $\hat{V}_{itest}^c = \hat{\tau}^2 Z d_{itest} d_{itest}' Z' + \hat{\sigma}^2 I_{J \times n}$

Test du type rapport de vraisemblance

En phase test, la classe d'appartenance de chaque nouvel essai $itest$ est déterminé par un test du type rapport de vraisemblance basé sur l'approche de Henderson *et al.*[?]. Sous l'hypothèse " y_{itest} appartient à la classe c ", la densité de probabilité du couple (y, b) au point (y_{itest}, b_{itest}) est donnée par

$$f_{(y,b)}^c(y_{itest}, b_{itest}) = f_{y|b=b_{itest}}^c(y_{itest}) f_b^c(b_{itest})$$

où $f_{y|b=b_{itest}}^c$ est la densité de la loi $\mathcal{N}(\mu^c + Z d_{itest} b_{itest}^c, \sigma_c^2)$.

Soit $l^c = \ln(f^c)$, la log-*quasi*-vraisemblance du modèle de la classe c , la statistique du test est obtenue par

$$S = l^1 - l^0 \quad \text{où } l^c \propto -\frac{1}{2\sigma^{2,c}} \|y_{itest} - \mu^c - Z_{itest} b_{itest}^c\|^2 - \frac{1}{2\tau^{2,c}} (b_{itest}^c)$$

en remplaçant les paramètres par leurs estimations et b_{itest} par sa prédiction $\widehat{b_{itest}^c}$.

La performance du test est évaluée par l'étude de la courbe ROC (Receiver Operating Characteristics) et de l'aire sous cette courbe (AUC : Area Under Curve).

3 Application à la détection d'ondes d'erreurs en EEG

L'approche présentée est appliquée à la détection d'ondes d'erreurs au cours d'une tâche cognitive (cf. Roger *et al.* (2010)). Sur un ensemble de six sujets, nous cherchons à classer deux

situations : le sujet s'est trompé au cours de la tâche (les essais *Erreur*) *versus* le sujet a répondu correctement au stimulus (les essais *Correct*). Les résultats sont comparés à ceux d'une AFD sur les électrodes utilisant la distance de Mahalanobis [1].

Les estimations des paramètres des modèles ont été obtenues à l'aide de l'algorithme Matlab *mixed.m* basé sur le système d'équations du modèle mixte d'Henderson [9].

Extraction de l'information pertinente

Pour chaque sujet, les données brutes sont de mêmes dimensions : les signaux EEG sont enregistrés à l'aide de 64 électrodes placées sur l'ensemble du scalp et sur un échantillon temporel de 128 instants. En phase d'apprentissage et en phase test, 30 essais de chaque classe *Erreur* et *Correct* sont tirés aléatoirement.

L'étape de pré-traitement par AFD sur les électrodes permet de réduire de manière significative la dimension spatiale : pour les six sujets, le nombre J de canaux est déterminé par le pourcentage d'inertie que l'on souhaite conserver. Dans cette application, ce taux d'inertie est fixé à 95%, et J varie de 2 à 5 selon les sujets.

Après application de la DWT sur les données temporelles, les deux premiers niveaux de décomposition, qui ne contiennent pas d'information pertinente pour la modélisation, sont supprimés, permettant de passer de 128 instants à 32 coefficients d'ondelettes.

La modélisation par le MLMG des coefficients en ondelettes permet d'extraire les caractéristiques propre à chaque classe de signaux, comme l'illustre la Figure 2:

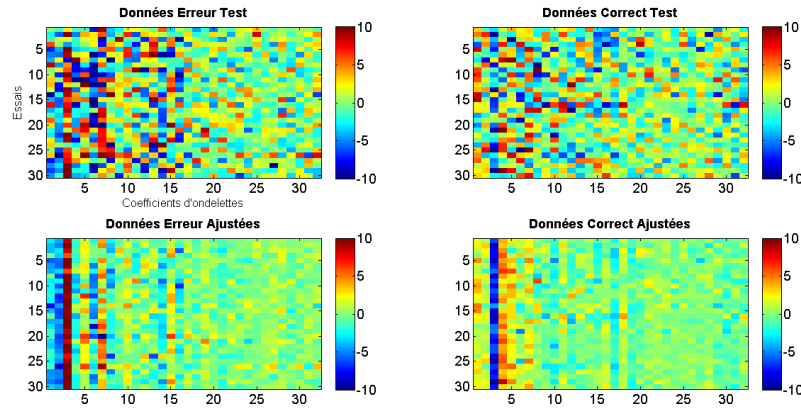


Figure 2: Pour le sujet 4, les données classées en phase test sont représentées sur le canal 1. En haut : les données test d'origine pour les *Erreurs* (à gauche) et les essais *Correct* (à droite). En bas : les données ajustées par le MLMG les essais *Erreur* et *Correct*

Performance du classifieur

Pour chaque sujet, les résultats obtenus par la procédure proposée sont comparés aux résultats obtenus par une AFD sur les électrodes. La représentation des courbes ROC, le calcul des AUC et leurs intervalles de confiance illustrent l'efficacité de notre méthode.

Pour quatre sujets (2-4-5-6), les résultats du classifieur basé sur le MLMG (*MEM* en bleu) sont meilleurs que ceux obtenus par l'AFD sur les électrodes (*AFD* en rouge). Pour les deux

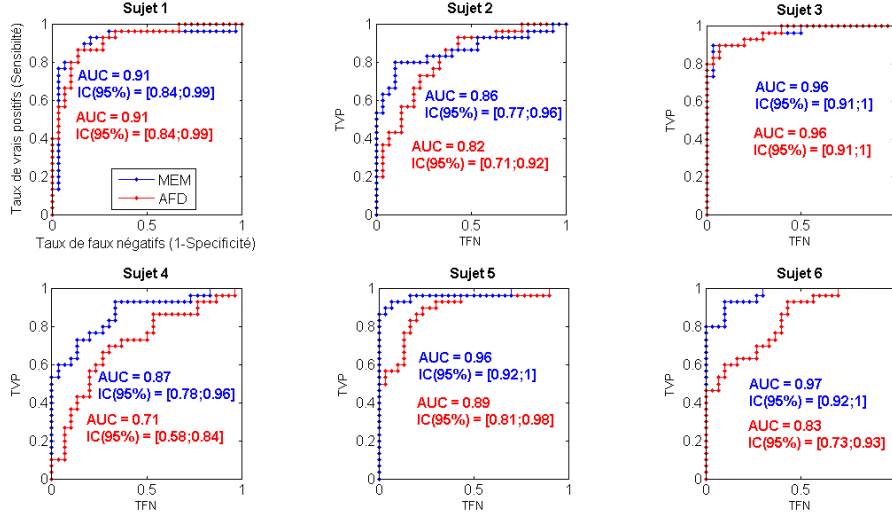


Figure 3: Pour chacun des six sujets étudiés sont représentés la courbe ROC du classifieur basé sur le MLMG (*MEM* en bleu) et celle obtenue par l’AFD sur les électrodes (*AFD* en rouge) et sont indiqués les AUC avec leur intervalles de confiance.

sujets restants (1-3), les résultats sont similaires.

L’apport de la procédure proposée réside non seulement dans une bonne classification de deux types de signaux EEG avec des échantillons restreints en phase d’apprentissage, mais aussi dans une meilleure compréhension des phénomènes neurophysiologiques étudiés. La modélisation permet d’extraire des caractéristiques discriminantes propres à chacune des classes.

4 Conclusion

Nous proposons une méthode de classification supervisée qui est à la fois descriptive et discriminante, basée sur la modélisation par le MLMG, pour des signaux fortement bruités tels que les signaux cérébraux. Les résultats obtenus sur les données EEG de six sujets au cours d’une tâche cognitive démontrent l’efficacité de la procédure. Ces résultats ont un grand intérêt applicatif comme par exemple pour les Interfaces Cerveau-Machine. La simplification du modèle proposé par rapport à l’approche de Huang *al* () et la réduction importante de la dimension temporelle sans perte d’information induite par l’introduction d’une DWT, permettent de réduire le coût de calcul de notre procédure. Ceci nous permet d’envisager la construction d’un modèle MLMG intégrant la dimension *individu* par la modélisation de la variabilité inter-sujets.

Bibliographie

- [2] Davidson, D.J. (2009), *Functional Mixed-Effect Models for Electrophysiological Responses*, Neurophysiologie, Vol. 41, No. 1, 71–87.
- [3] Fazli, S., Danoczy, M., Schelldorfer, J., Müller, K.R. (2011), *l1-penalized mixed-effects models for high dimensional data with application to BCI*, NeuroImage, 56, 2100–2108.
- [4] Gutiérrez, D., Escalona-Vargas, D.I. (2010), *EEG data classification through signal spatial re-distribution and optimized linear discriminants*, Computer Method and Programs in Biomedicine, 97, 39–47.
- Henderson, C. R., Kempthorne, O., Searle, S. R., von Krosigk, C. M. (1959). The Estimation of Environmental and Genetic Trends from Records Subject to Culling. Biometrics 15 (2), 192-218.
- [5] Huang, Y., Erdogmus, K., Pavel, M., Mathan, S. (2008), *Mixed Effects Models for EEG Evoked Response Detection*, IEEE Workshop on Machine Learning for Signal Processing, Cancun, Mexico.
- [6] Huang, Y., Erdogmus, D., Hild, K., Pavel, M., Mathan, S. (2009), *Mixed Effects Models for Single-Trial ERP Detection in Noninvasive Brain Computer Interface Design*, E-Book Preprint Bentham Science Publishers.
- [8] Roger, C., Bénar, C.G., Vidal, F., Hasbroucq, T., Burle, B. (2010), *Rostral Cingulate Zone and correct response monitoring : ICA and source localization evidences for the unicity of correct- and error-negativities*, NeuroImage, 51, 391–403.
- [9] Vitkovsky, V. (2002), *MATLAB algorithm mixed.m for solving Henderson's mixed model equations*, Technical Report, Institute of Measurement Science, Bratislava.